TECHNICKÁ UNIVERZITA V LIBERCI

Ekonomická fakulta

BAKALÁŘSKÁ PRÁCE

2012 **Petr Kusák**

TECHNICKÁ UNIVERZITA V LIBERCI

Ekonomická fakulta

Studijní program: 6208 Ekonomika a management

Studijní obor: Podniková ekonomika

Klasické a adaptivní metody pro vyrovnávání časových řad ekonomických ukazatelů.

Classic and adaptive methods for smoothing time series of economic data.

Vedoucí práce: Mgr. Jiří, Rozkovec, Katedra statistiky

Konzultant: Ing. Vladimíra, Hovorková Valentová PhD., Katedra statistiky

Počet stran: 62 Počet příloh: 1

Datum odevzdání: 04. 01. 2012

Prohlášení

Byl jsem seznámen s tím, že na mojí bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu TUL.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím diplomové práce a konzultantem.

Datum: Podpis: ....

Poděkování

Rád bych na tomto místě vyjádřil poděkování všem, kteří mi s přípravou mé bakalářské práce pomáhali, protože bez nich by tato práce nemohla vzniknout.

Zejména paní Ing. Hovorkové-Valentové PhD., vedoucí bakalářské práce, za odbornou pomoc, rady a připomínky k bakalářské práci.

Resumé

Moje bakalářská práce se zaměřuje na popis a analýzu časových řad pomocí klasických a adaptivních metod a zdůvodňuje jejich význam pro praxi.

Celou práci jsem koncipoval také jako návod pro studenty statistiky, kteří mohou mít se statistikou problémy, ale potřebují na toto téma vypracovat například semestrální práci.

Summary

My bachelor work is focused on the description and analysis of the time series with classic and adaptive methods and gives reasons of its importance for the practical purposes.

The whole paper was also designed as a comprehensive manual for those students who are not familiar with statistics but are in need to be able to piece together some seminar work on this subject.

Klíčová slova

Analýza

Časové řady

Index determinace

Interval spolehlivosti

Klouzavé průměry

Přípustná chyba odhadu

Součet čtverců chyb

Střední čtvercová chyba odhadu

Transformovaná časová proměnná

Vyrovnávací konstanta

Vyrovnávání

Zpracování a analýza dat

Key words

Analysis

Time Series

Determination Index (I^{2})

Confidelity Interval

Moving Averages

Acceptable Estimation Error

Sum Of Squared Errors (S.S.E.)

Mean Squared Error (M.S.E.)

Transformed Time Variable

Smoothing Constant

Smoothing

Data processing and analysis

Obsah

1 Úvod

Tato bakalářská práce se zabývá vyrovnáváním časových řad pomocí klasických a adaptivních metod a jejich srovnáním. Z klasických metod se zaměřuje na metodu nejmenších čtverců u lineární, parabolické a exponenciální trendové funkce. Z adaptivních metod je ukázána metoda klouzavých průměrů a Brownova metoda jednoduchého a dvojitého exponenciálního vyrovnávání.

Každý postup je podrobně vysvětlen a doplněn praktickým příkladem a grafem konkrétní časové řady.

Na závěr je uveden souhrnný příklad, který srovnává vyrovnávání konkrétní časové řady lineární trendovou funkcí a Brownovou metodou dvojitého exponenciálního vyrovnávání.

2 Základní pojmy

2.1 Definice časové řady

Časová řada je posloupnost hodnot určitého statistického znaku (ukazatele) uspořádaných z hlediska času ve směru od minulosti k přítomnosti. Ukazatel musí být věcně a prostorově shodně vymezen po celé sledované období.

2.2 Historie analýzy časových řad

První významný krok v analýze časových řad byl učiněn Yuleho autoregresním modelem a Slutského modelem klouzavých průměrů v roce 1927. Yule popsal vlastnosti autoregresních modelů, zavedl parciální autokorelace a odhadl autoregresní model nízkého řádu pomocí metody nejmenších čtverců na základě konkrétní časové řady. Wold byl v roce 1938 prvním, kdo odhadl model klouzavých průměrů na základě reálných dat.

Ve čtyřicátých letech byl učiněn pokrok v odhadovacích procedurách. Byla odvozena asymptotická teorie pro odhady parametrů v autoregresních modelech. Současně byla odvozena i asymptotická teorie pro výběrové autokorelace.

V roce 1952 Whittle, pravděpodobně jako první, pomocí těchto modelů zachycoval sezónnost. V padesátých a šedesátých letech bylo o této problematice napsáno mnoho prací, jejichž vyústěním byla v roce 1970 vydaná dnes již klasická Boxova a Jenkinsova práce „Time Series Analysis: Forecasting and Control“. Díky rozvoji počítačů a počítačového softwaru došlo v následujících letech k velkému rozšíření modelů ARIMA, jenž používají k odstranění nesezónní a sezónní nestacionarity v časových řadách nesezónní a sezónní diference. Ačkoliv byly tyto transformace známy již relativně dlouhou dobu, právě Box a Jenkins je zpopularizovali.

Poněkud jiným přístupem pro modelování časových řad je spektrální analýza. Spektrální analýza byla známa již před prvními pokusy o modelováním časových řad v tzv.

časové doméně. V roce 1898 prezentoval Schuster svůj periodogram, který se později používal především ke zkoumání periodicity časových řad. V roce 1929 navrhl Fisher test pro zjišťování periodicity v časových řadách. S vývojem počítačového softwaru se spektrální analýza značně zpopularizovala, přesto se ale při modelování ekonomických časových řad obecně nerozšířila.

Zdroj: [5]

2.3 Druhy časových řad

Rozdělení časových řad podle časového hlediska

- Časová řada okamžiková: příslušný ukazatel udává, stav sledovaného ukazatele v daném časovém okamžiku (např. počet obyvatelstva k určitému dni).
- Časová řada intervalová: příslušný ukazatel udává, kolik jevů vzniklo či zaniklo v určitém časovém intervalu (např. počet sňatků během roku). Nejsou-li jednotlivé časové intervaly ekvidistantní (stejné délky), musí se provést očištění časové řady od důsledků kalendářních variací.

Rozdělení časových řad podle periodicity

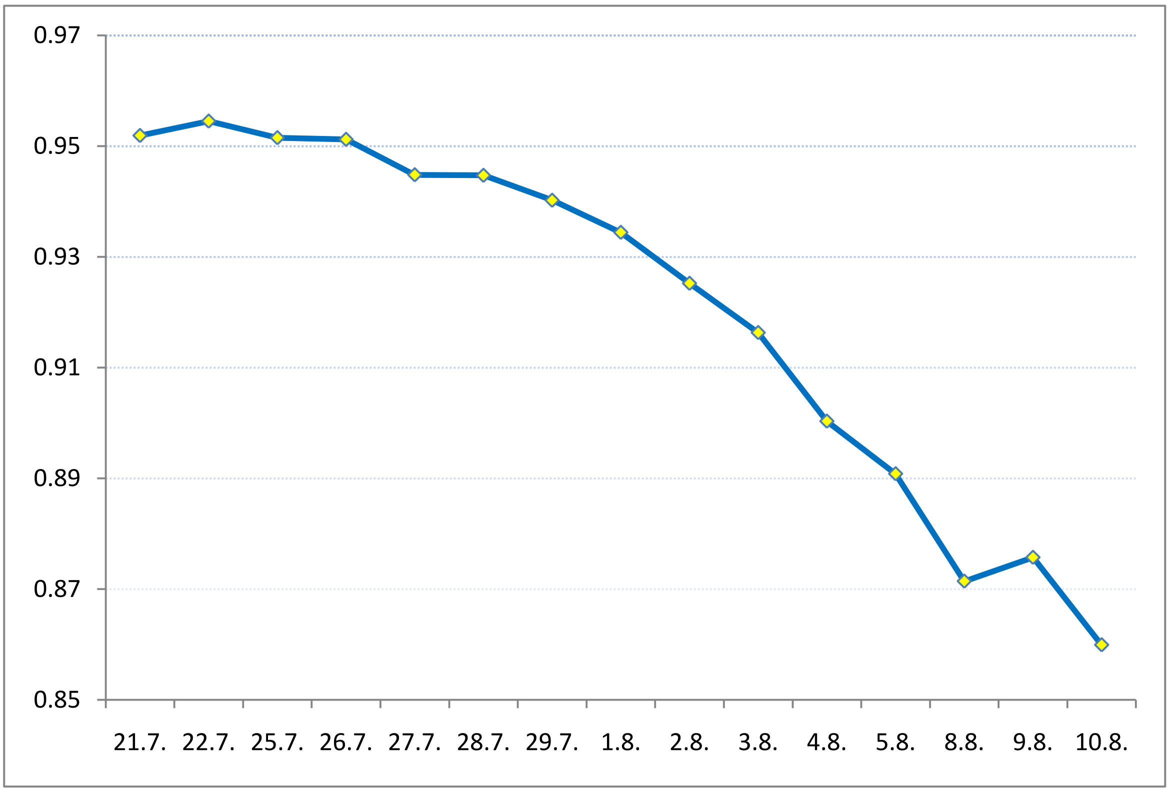
- Časové řady dlouhodobé. Jejich periodicita je jeden rok a více. Aplikují se jiné postupy než u časových řad krátkodobých.
- Časové řady krátkodobé. S periodicitou kratší než jeden rok. Čtvrtletní, měsíční, týdenní, atd.

Rozdělení časových řad podle způsobu vyjádření ukazatelů

- Časové řady naturálních ukazatelů. Hodnoty jsou vyjádřeny v naturálních jednotkách. Například objem produkce v tunách.
- Časové řady peněžních ukazatelů. Hodnoty jsou vyjádřeny v penězích. Například výnosy investice za jednotlivá období. Nutno zajistit srovnatelnost v čase, existuje například nebezpečí změny cenové hladiny.

2.4 Grafické znázornění časové řady

Okamžikovou časovou řadu graficky znázorňujeme pomocí spojnicového diagramu. Na vodorovnou osu vynášíme časové okamžiky $t_1, \ldots, t_n$, na svislou osu odpovídající hodnoty $y_1, \ldots, y_n$. Dvojice bodů (t_i, y_i), $i = 1, \ldots, n$ následně spojíme úsečkami.

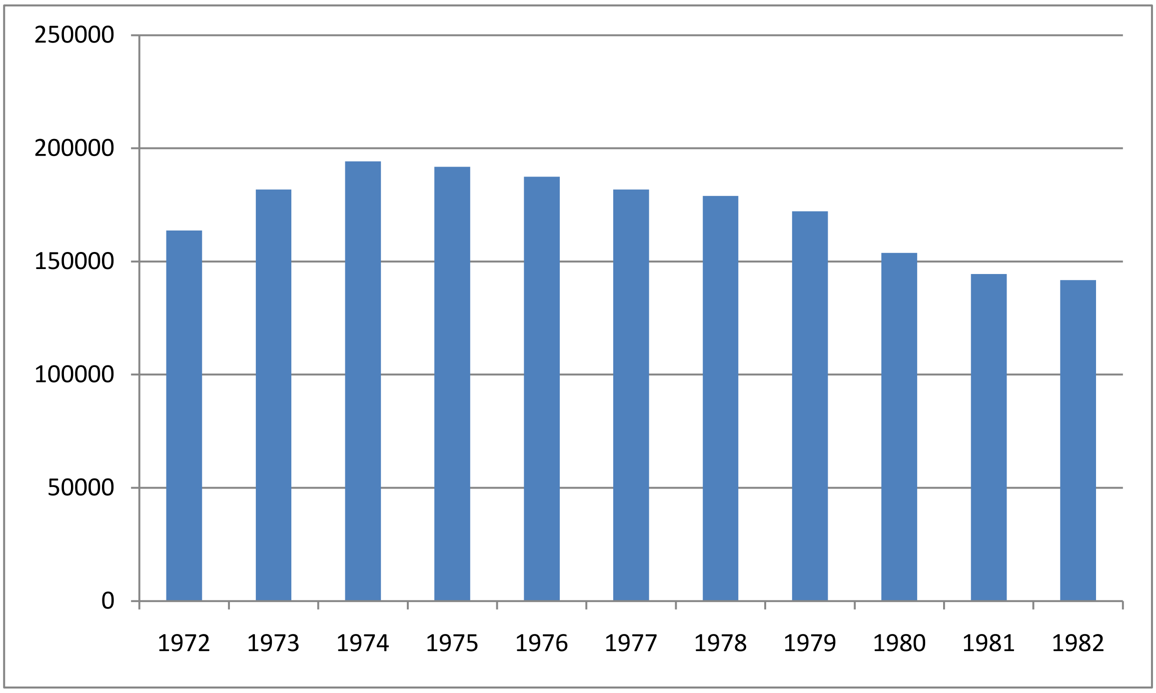


Obrázek 1: Grafické znázornění okamžikové časové řady

Zdroj: http://www.pioneer.cz/Fond/HistorickeCenyVysledky.asp

Hodnoty cen podílového listu pro **Pioneer - dynamický fond** za období od **21.7.2011** do **10.8.2011**

Intervalovou časovou řadu nejčastěji znázorňujeme sloupkovým diagramem. Je to soustava obdélníků, kde šířka obdélníku je rovna délce intervalu a výška odpovídá hodnotě ukazatele v daném intervalu. Ke znázornění intervalové časové řady lze použít i spojnicový diagram, přičemž na vodorovnou osu vynášíme středy příslušných intervalů.



Obrázek 2: Grafické znázornění intervalové časové řady

Zdroj: Český statistický úřad:
Počet narozených dětí v letech 1972-1982 v ČSSR

2.5 Popisné charakteristiky časových řad

V dalším textu budou vysvětleny nejčastěji používané popisné charakteristiky časových řad. (viz [7])

Průměr okamžikové časové řady

Jsou-li všechny intervaly mezi jednotlivými hodnotami časové řady stejně dlouhé, vypočteme prostý chronologický průměr okamžikové časové řady:

$$y = \frac{\frac{y_1}{2} + y_2 + \ldots + y_{n-1} + \frac{y_n}{2}}{n-1}$$

Nemají-li intervaly stejnou délku, použijeme vážený chronologický průměr okamžikové časové řady:

$$y = \frac{\sum_{i=1}^{n-1} y_i \cdot d_i}{\sum_{i=1}^{n-1} d_i},$$

kde d_i je vzdálenost mezi dvěma sousedními časovými okamžiky (délka intervalu).

Průměr intervalové časové řady

Průměr intervalové časové řady počítáme podle vzorce:

$$y = \frac{1}{n}\sum_{i=1}^{n} y_i$$

jako aritmetický průměr jednotlivých hodnot.

2.6 Dynamické charakteristiky časových řad

Absolutní přírůstky

Jedná se o první, druhé a další diference, tj. rozdíly mezi sousedními hodnotami časové řady.

1. diference: $\Delta y_i = y_i - y_{i-1}$, $i = 2, ..., n$

2. diference: $\Delta^{(2)} y_i = \Delta y_i - \Delta y_{i-1} = y_i - 2y_{i-1} + y_{i-2}$, $i = 3, ..., n$ atd.,

obecně potom diference řádu k: $\Delta^{(k)} y_i = \Delta^{(k-1)} y_i - \Delta^{(k-1)} y_{i-1}$, $i = k+1, ..., n$.

Diferencování má velký význam při odhadu trendu časové řady regresními metodami.

Průměrný absolutní přírůstek – aritmetický průměr prvních diferencí

$$\overline{\Delta} = \frac{y_n - y_1}{n-1}$$

Relativní přírůstek

$$\delta_i = \frac{y_i - y_{i-1}}{y_{i-1}} = \frac{\Delta y_i}{y_{i-1}}, \qquad i = 2,\ 3,\ ...,\ n$$

Relativní přírůstek po vynásobení 100 udává, o kolik procent se změnila hodnota sledovaného ukazatele v čase t_i oproti t_{i-1}.

Koeficient růstu – udává kolikrát vzrostla hodnota časové řady v časovém okamžiku t_i oproti období předcházejícímu.

$$k_i = \frac{y_i}{y_{i-1}}, i = 2,3,...n$$

Koeficient růstu po vynásobení 100 udává, na kolik procent hodnoty v čase t_{i-1} vzrostla či poklesla hodnota ukazatele v čase t_i.

Průměrný koeficient růstu – geometrický průměr jednotlivých koeficientů růstu

$$\bar{k} = \sqrt[n-1]{k_2 \cdot k_3 \cdot \ldots \cdot k_n} = \sqrt[n-1]{\frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdot \ \ldots \ \cdot \frac{y_n}{y_{n-1}}} = \sqrt[n-1]{\frac{y_n}{y_1}}$$

3 Analýza časových řad

Klasická analýza časových řad vychází z předpokladu, že se časová řada skládá ze tří složek - **trendové**, **periodické (cyklické**, **sezónní)** a **reziduální**.

- **Trendovou složkou** se rozumí dlouhodobá tendence ve vývoji časové řady.
- **Periodickou složku** můžeme rozdělit na dva typy:
    - **Cyklickou složkou** se chápe dlouhodobé pravidelné či nepravidelné kolísání okolo trendu, někdy se také hovoří o fluktuacích okolo trendu. Délka jednotlivých cyklů, jakož i jejich intenzita je proměnlivá a je delší než 1 rok.
    - **Sezónní složka** popisuje periodické změny v časové řadě, které se odehrávají během jednoho kalendářního roku a každý rok se pravidelně opakují. Příčiny způsobující sezónnost v časové řadě jsou např. změny ročních období a s nimi spojené změny počasí a délky slunečního svitu, různé institucionálně zakotvené lidské zvyky - prázdniny, náboženské svátky, různá výročí atd. Období je delší než 1 rok.
- **Reziduální složka** zbývá v časové řadě po odstranění trendové, cyklické a sezónní složky. Je tvořena mimo jiné náhodnými pohyby v průběhu časové řady.

Z výše uvedeného rozdělení časových řad je patrné, že všechny zmíněné složky jsou obsaženy pouze v časových řadách krátkodobých (dlouhodobé časové řady nemohou obsahovat sezónní složku).

Zdroj: [3]

3.1 Trendová analýza.

Existuje několik metod trendové analýzy. V této práci se zaměřím na nejčastěji používanou metodu nejmenších čtverců.

Pro demonstraci celého postupu použijeme tři trendové funkce, a to lineární, parabolickou a exponenciální.

Podstatou metody nejmenších čtverců je nalezení funkce $f(t) = Y_t$, která bude popisovat závislost sledované veličiny y_t na čase t a která má takové parametry $\beta_0, \beta_1, \ldots, \beta_n$, aby součet

$$\sum_{t=1}^{n}(y_t - Y_t)^2 \text{ byl minimální,}$$

kde $t = 1, 2, ..., n$ je časová proměnná,

Y_t je jsou vyrovnané hodnoty časové řady,

a y_t jsou pozorované hodnoty časové řady.

tj. součet druhých mocnin rozdílu mezi skutečnými a vypočítanými hodnotami dané časové řady je nejmenší. Tuto funkci odhadneme pomocí čísel b_0, b_1 .

Rozdíl $e_t = y_t - Y_t$ se nazývá *reziduum* a součet $\sum_{t=1}^{n}(y_t - Y_t)^2$ potom *reziduálním součtem čtverců*. Dále jako SSE (Sum of Squared Errors).

3.1.1 Lineární trendová funkce

U lineární trendové funkce platí:

$$Y_t = \beta_0 + \beta_1 \cdot t\,,$$

kde t je 1, 2, ..., n a

β_0, β_1 jsou neznámé parametry, jejichž odhady značíme b_0 a b_1.

Potom součet je roven $\quad S(b_0, b_1) = \sum_{t=1}^{n} \left(y_t - b_0 - b_1 \cdot t\right)^2$

Tento součet má být minimální, proto spočteme obě parciální derivace (podle b_0 a b_1) a tyto položíme rovny nule. Tak určíme oba odhady b_0 a b_1.

Derivace podle b_0 :

$$\frac{dS}{db_0} = \sum 2 \cdot \left(y_t - b_0 - b_1 \cdot t\right) \cdot \left(-1\right) = \sum -2 \cdot \left(y_t - b_0 - b_1 \cdot t\right) =$$

$$= -\sum 2y_t + \sum 2b_0 + \sum 2b_1 t = -\sum 2y_t + 2nb_0 + \sum 2b_1 t =$$

$$\boxed{= -2 \cdot \sum y_t + 2nb_0 + 2b_1 \cdot \sum t}$$

Derivace podle b_1 :

$$\frac{dS}{db_1} = \sum 2 \cdot \left(y_t - b_0 - b_1 \cdot t\right) \cdot \left(-t\right) = \sum -2t \cdot \left(y_t - b_0 - b_1 \cdot t\right) =$$

$$= -\sum 2y_t t + \sum 2b_0 t + \sum 2b_1 t^2 =$$

$$\boxed{= -2 \cdot \sum y_t t + 2b_0 \cdot \sum t + 2b_1 \cdot \sum t^2}$$

Obě tyto derivace položíme rovny nule, čímž dostaneme tzv. *soustavu normálních rovnic*, ze kterých odhadneme parametry b_0, b_1 :

$$-2 \cdot \sum y_t t + 2nb_0 + 2b_1 \cdot \sum t = 0$$

$$-2 \cdot \sum y_t t + 2b_0 \cdot \sum t + 2b_1 \cdot \sum t^2 = 0$$

$$-\sum y_t + nb_0 + b_1 \cdot \sum t = 0$$

$$-\sum y_t t + b_0 \cdot \sum t + b_1 \cdot \sum t^2 = 0$$

Řešením této soustavy dostaneme následně vzorce pro odhady parametrů b_0, b_1.

Z první rovnice vyjádříme b_0

$$nb_0 = \sum y_t - b_i \sum t$$

$$\boxed{b_0 = \frac{\sum y_t - b_1 \sum t}{n}} \qquad (1.)$$

A dosadíme do druhé:

$$-\sum y_t t + \frac{\sum y_t \cdot \sum t - b_1 \left(\sum t\right)^2}{n} + b_1 \cdot \sum t^2 = 0 \qquad / \cdot n$$

$$-n \sum y_t t + \sum y_t \cdot \sum t - b_1 \left(\sum t\right)^2 + nb_1 \sum t^2 = 0$$

$$nb_1 \sum t^2 - b_1 \left(\sum t\right)^2 = n \sum y_t t - \sum y_t \cdot \sum t$$

$$b_1 \left(n \sum t^2 - \left(\sum t\right)^2\right) = n \sum y_t t - \sum y_t \cdot \sum t$$

$$\boxed{b_1 = \frac{n \sum y_t t - \sum y_t \cdot \sum t}{n \sum t^2 - \left(\sum t\right)^2}} \qquad (2.)$$

Vydělíme-li čitatele i jmenovatele výrazu (2.) výrazem n^2 dostaneme:

$$b_1 = \frac{\frac{\sum y_t t}{n} - \frac{\sum y_t}{n} \cdot \frac{\sum t}{n}}{\frac{\sum t^2}{n} - \left(\frac{\sum t}{n}\right)^2},$$

To je kovariance proměnných y a t děleno rozptylem proměnné t.

Takže $$\boxed{b_1 = \frac{\operatorname{cov}(y,t)}{\sigma^2{}_t}} \qquad (3.)$$

Výraz (1.) lze stejným způsobem upravit na:

$$b_0 = \frac{\sum y_t}{n} - b_1 \cdot \frac{\sum t}{n}$$

$$\boxed{b_0 = \bar{y} - b_1 \cdot \bar{t}} \qquad (4.)$$

To je průměr hodnot časové řady mínus průměr $t_1,...,t_n$ vynásobený koeficientem b_1.

Pomocí vzorců (3.) a (4.) můžeme již určit hledané koeficienty b_0 a b_1.

Nebo můžeme výraz (2.) dosadit do první rovnice:

$$nb_0 = \sum y_t - \sum t \cdot \frac{n\sum y_t t - \sum y_t \cdot \sum t}{n\sum t^2 - \left(\sum t\right)^2}$$

$$nb_0 = \frac{n\sum y_t \cdot \sum t^2 - \sum y_t \cdot \left(\sum t\right)^2 - n\sum y_t t \cdot \sum t + \sum y_t \cdot \left(\sum t\right)^2}{n\sum t^2 - \left(\sum t\right)^2}$$

$$nb_0 = \frac{n\sum y_t \cdot \sum t^2 - n\sum y_t t \cdot \sum t}{n\sum t^2 - \left(\sum t\right)^2}$$

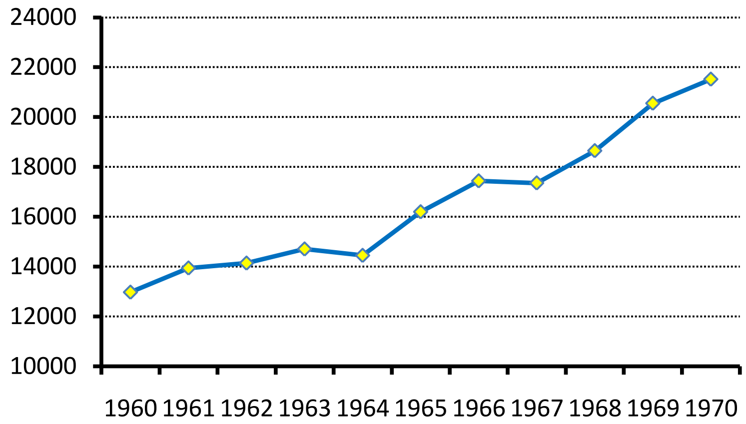
$$\boxed{b_0 = \frac{\sum y_t \cdot \sum t^2 - \sum y_t t \cdot \sum t}{n\sum t^2 - \left(\sum t\right)^2}} \qquad (5.)$$

Potom pomocí výrazů (2.) a (5.) můžeme přímo vypočítat hledané koeficienty b_0 a b_1.

Příklad 1.

Hodnoty o počtech rozvodů v letech 1960 – 1970 z tabulky:
http://www.czso.cz/csu/edicniplan.nsf/t/93003E0F91/$File/tab1.xls vyrovnáme pomocí lineární trendové funkce a odhadneme hodnoty na následující tři roky.

rok	počet rozvodů
1960	12970
1961	13939
1962	14137
1963	14703
1964	14446
1965	16196
1966	17435
1967	17352
1968	18647
1969	20550
1970	21516



Řešení:

Tabulka 1.

rok	t	y_t	y_t .t	t^2	Y_t
1960	1	12970	12970	1	12399,59
1961	2	13939	27878	4	13226,78
1962	3	14137	42411	9	14053,97
1963	4	14703	58812	16	14881,16
1964	5	14446	72230	25	15708,35
1965	6	16196	97176	36	16535,55
1966	7	17435	122045	49	17362,74
1967	8	17352	138816	64	18189,93
1968	9	18647	167823	81	19017,12
1969	10	20550	205500	100	19844,31
1970	11	21516	236676	121	20671,50
Σ	**66**	**181891**	**1182337**	**506**	**181891**
∅	**6**	**16535.55**	**107485**	**46**	**16535.55**

Zdroj: vlastní výpočet

$$b_1 = \frac{\text{cov}(y,t)}{\sigma^2(t)} = \frac{107485{,}2 - 6 \cdot 16535{,}55}{46 - 6^2} = \frac{8271{,}909}{10} = 827{,}19$$

$$b_0 = y - b_1 \cdot t = 16535{,}55 - 827{,}19 \cdot 6 = 11572{,}41$$

Výsledná lineární trendová funkce má tedy tvar:

$$\boxed{Y_t = 11572{,}4 + 827{,}19 \cdot t}$$

$t = 1, ..., 11$

Jednotlivé vypočítané hodnoty jsou v posledním sloupci tabulky 1.

Použijeme-li k výpočtu koeficientů vzorce (2.) a (5.) dostaneme stejný výsledek.

$$b_1 = \frac{n\sum y_t t - \sum y_t \cdot \sum t}{n\sum t^2 - \left(\sum t\right)^2} = \frac{11 \cdot 1182337 - 181891 \cdot 66}{11 \cdot 506 - 66^2} = 827{,}19$$

$$b_0 = \frac{\sum y_t \cdot \sum t^2 - \sum y_t t \cdot \sum t}{n\sum t^2 - \left(\sum t\right)^2} = \frac{181891 \cdot 506 - 1182337 \cdot 66}{11 \cdot 506 - 66^2} = 11572{,}41$$

Pomocí nalezené trendové funkce můžeme potom spočítat předpovědi na další roky
(za předpokladu, že se trend nezmění):

$$Y_{1971} = 11572{,}4 + 827{,}19 \cdot 12 = 21499$$

$$Y_{1972} = 11572{,}4 + 827{,}19 \cdot 13 = 22326$$

$$Y_{1973} = 11572{,}4 + 0827{,}19 \cdot 14 = 23153$$

Skutečné hodnoty přitom byly:

$y_{1971} = 23616$ $\quad$ $y_{1972} = 22392$ $\quad$ $y_{1973} = 25271$

Kromě právě uvedených bodových předpovědí můžeme se stanovenou pravděpodobností určit intervalové předpovědi. To znamená, že vypočteme interval, ve kterém se odhadovaná hodnota bude nacházet s danou pravděpodobností.

Pro stanovení intervalu použijeme vzorec

$\hat{Y}_t - t_{1-\alpha/2}[n-2]\cdot s\cdot g_p < Y_t < \hat{Y}_t + t_{1-\alpha/2}[n-2]\cdot s\cdot g_p$, kde

$t_{1-\alpha/2}[n-2]$ je (1-α/2).100% kvantil t-rozdělení o n-2 stupních volnosti,

přičemž součin $t_{1-\alpha/2}[n-2].s.g_P$ je tzv. *přípustná chyba odhadu*,

$$s = \sqrt{\frac{\sum y_t^2 - \sum Y_t^2}{n-2}}, \tag{6.}$$

$$g_p = \sqrt{1 + \frac{1}{n} + \frac{(P-\bar{t})^2}{\sum t^2 - n\bar{t}^2}}, \text{ kde P je index odhadovaného roku.} \tag{7.}$$

v našem případě 12, 13 a 14.

Budeme chtít předpovědní interval s 90% spolehlivostí.

$n - 2 = 9$ $\alpha = 0{,}1$.

V tabulce kvantilů t-rozdělení najdeme $t_{0,95}[9] = 1{,}833$. Hodnoty s a g_p získáme z následující tabulky:

Tabulka 2.

t	**$y_t^2 . 10^6$**	**$Y_t^2 . 10^6$**
1	168,22	153,75
2	194,29	174,95
3	199,85	197,51
4	216,18	221,45
5	208,69	246,75
6	262,31	273,42
7	303,98	301,46
8	301,09	330,87
9	347,71	361,65
10	422,31	393,79
11	462,94	427,31
součet	**3 087,57**	**3 082,93**

Pomocí vzorců (6.) a (7.) vypočítáme

$s = 717{,}68$ a $g_{12} = 1.19$, $g_{13} = 1.24$, $g_{14} = 1.29$

Přípustná chyba odhadu pro další roky bude:

Δ_{1971} = 1,833 . 717,68 . 1,19 = 1565
Δ_{1972} = 1,833 . 717,68 . 1,24 = 1631
Δ_{1973} = 1,833 . 717,68 . 1,29 = 1697

Předpovědní intervaly potom budou:
1971: 21499 - 1565 < y_{1971} < 21499 +1565, tj. (19934;23064)
1972: 22326 – 1631 < y_{1972} < 22326 + 1631, tj. (20695;23957)
1973: 23153 – 1697 < y_{1973} < 23153 + 1697, tj. (21456;24850)

Porovnáme-li vypočtené intervaly se skutečnými hodnotami, zjistíme, že ve dvou případech (1971 a 1973) leží skutečná hodnota mimo interval.

Můžeme chtít větší hladinu spolehlivost (99%), která s sebou ovšem nese i rozšíření intervalu spolehlivosti.

(n – 2 = 9 α = 0,01 $t_{0,995}[9]$ = 3,25)

Další postup je analogický pomocí tabulky 2. a vzorců (6.) a (7.)

s = 717,68 a g_{12} = 1.19, g_{13} = 1.24, g_{14} = 1.29 zůstávají stejné

Přípustná chyba odhadu pro další roky bude:

Δ_{1971} = 3,25 . 717,68 . 1,19 = 2775
Δ_{1972} = 3,25 . 717,68 . 1,24 = 2892
Δ_{1973} = 3,25 . 717,68 . 1,29 = 3008

Předpovědní intervaly pro 99% spolehlivost :
1971: 21499 – 2775 < y_{1971} < 21499 + 2775, tj. (18724;24274)
1972: 22326 – 2892 < y_{1972} < 22326 + 2892, tj. (19434;25218)
1973: 23153 – 3008 < y_{1973} < 23153 + 3008, tj. (20145;26161)

Při spolehlivosti předpovědi 99% dostaneme širší intervaly předpovědi, které již obsahují skutečné hodnoty.

K posouzení vhodnosti trendové funkce existuje řada ukazatelů, z nichž nejpoužívanější je M.S.E. (Mean Squared Error), tzv. *střední čtvercová chyba odhadu*.

$$M.S.E. = \frac{\sum_{t=1}^{n} (y_t - Y_t)^2}{n}$$

Pro konkrétní časovou řadu je nejvhodnější ta trendová funkce, pro kterou vyjde střední čtvercová chyba nejmenší. Nutný je stejný počet parametrů u porovnávaných modelů.

Kromě toho můžeme určit *index determinace I^2*, abychom se ujistili, na kolik procent je daná lineární funkce vhodná k vyrovnání dané řady.

Tabulka 3.

rok	t	y_t	$y_t .t$	t^2	Y_t	S_Y	S_T	S_R
1960	1	12970	12970	1	12399.59	12713114	17106120	325366.5
1961	2	13939	27878	4	13226.78	6742048	10947917	507254.7
1962	3	14137	42411	9	14053.97	5753020	6158203	6893.528
1963	4	14703	58812	16	14881.16	3358223	2736979	31742.28
1964	5	14446	72230	25	15708.35	4366200	684244.8	1593539
1965	6	16196	97176	36	16535.55	115291.1	0	115291.1
1966	7	17435	122045	49	17362.74	809018.5	684244.8	5222.033
1967	8	17352	138816	64	18189.93	666598	2736979	702122.1
1968	9	18647	167823	81	19017.12	4458240	6158203	136987.5
1969	10	20550	205500	100	19844.31	16115845	10947917	497999.7
1970	11	21516	236676	121	20671.5	24804927	17106120	713180.3
součet	**66**	**181891**	**1182337**	**506**	**181891**	**79902527**	**75266928**	**4635599**

K určení indexu determinace použijeme hodnoty z posledních tří sloupců tabulky 3.

$S_Y = \sum (y_t - y)^2$ celkový součet čtverců

$S_T = \sum (Y_t - y)^2$ teoretický součet čtverců

$S_R = \sum (y_t - Y_t)^2$ reziduální součet čtverců

Index determinace se rovná $I^2 = \frac{S_Y}{S_T}$, přičemž S_Y se musí rovnat $S_T + S_R$,

S_T je součet čtverců odchylek hodnot časové řady od jejich aritmetického průměru a S_Y je součet čtverců odchylek vypočtených hodnot od aritmetického průměru.

$$S_Y = \sum (y_t - y)^2 = 79902527$$

$$S_T = \sum (Y_t - y)^2 = 75266928$$

Index determinace se potom rovná $I^2 = \dfrac{75266928}{79902527} = 0{,}942$

Tato hodnota znamená, že 94,2% variability sledovaného ukazatele lze vysvětlit pomocí lineární trendové funkce.

Celý výpočet lze značně zjednodušit, použijeme-li místo proměnné t proměnnou t', tj. tzv. transformovanou časovou proměnnou, pro kterou platí:

1) $\sum t'^n = 0$ pro lichá n
2) t' je aritmetická posloupnost

$$-\sum y_t + nb_0 + b_1 \cdot \sum t' = 0$$

$$-\sum y_t t' + b_0 \cdot \sum t' + b_1 \cdot \sum t'^2 = 0$$

Protože $\sum t' = 0$, původní rovnice se nám zjednoduší na

$$-\sum y_t + nb_0 = 0$$

$$-\sum y_t t' + b_1 \cdot \sum t'^2 = 0$$

Odsud

$$b_0 = \frac{\sum y_t}{n} = y \quad \text{a} \tag{8.}$$

$$b_1 = \frac{\sum y_t t'}{\sum t'^2} \tag{9.}$$

Tabulka 4.

rok	t'	y_t	$y_t \cdot t'$	t'^2	Y_t
1960	-5	12 970	-64 850	25	12 399,59
1961	-4	13 939	-55 756	16	13 226,78
1962	-3	14 137	-42 411	9	14 053,97
1963	-2	14 703	-29 406	4	14 881,16
1964	-1	14 446	-14 446	1	15 708,35
1965	0	16 196	0	0	16 535,55
1966	1	17 435	17 435	1	17 362,74
1967	2	17 352	34 704	4	18 189,93
1968	3	18 647	55 941	9	19 017,12
1969	4	20 550	82 200	16	19 844,31
1970	5	21 516	107 580	25	20 671,50
součet	**0**	**181 891**	**90 991**	**110**	**181 891**
průměr	**0**	**16 535,55**	**8271,90**	**10**	**16 535,55**

b_0 je průměr hodnot časové řady, tedy 16535,55

b_1 = 90991/110 = 827,19

Výsledná lineární trendová funkce má potom tvar:

$$\boxed{Y_t = 16535{,}55 + 827{,}19 \cdot t'}$$

Jak je zřejmé z posledního sloupce tabulky 4., přestože dostaneme jiný koeficient b_0, vypočítané hodnoty jsou úplně stejné. Stejně tak předpovědi na další roky. Při této transformaci, při lichém počtu empirických hodnot, zůstává odhad b_1 stejný, ale změní se b_0. Při sudém počtu hodnot bude odhad b_1 poloviční, b_0 se změní úplně.

$$Y_{1971} = 16535{,}55 + 827{,}19 \cdot 6 = 21499$$

$$Y_{1972} = 16535{,}55 + 827{,}19 \cdot 7 = 22326$$

$$Y_{1973} = 16535{,}55 + 827{,}19 \cdot 8 = 23153$$

3.1.2 Parabolický trend

Parabolická funkce má tvar:

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 \quad \text{, kde } t = 1, 2, ..., n$$

Odhad parametrů této funkce provedeme pomocí koeficientů b_0, b_1 a b_2. Reziduální součet čtverců je potom

$$S(b_0, b_1, b_2) = \sum_{t=1}^{n} \left(y_t - b_0 - b_1 t - b_2 t^2 \right)^2 \quad \rightarrow \text{ MIN.}$$

Opět provedeme parciální derivace. Tentokrát ale budou tři, podle b_0, b_1 a b_2.

Stejnou metodou jako u lineární trendové funkce spočteme odhady. Použijeme transformovanou časovou proměnnou. Všechny tři parciální derivace položíme rovny nule a dostaneme soustavu tří rovnic se třemi neznámými b_0, b_1 a b_2. Po vyřešení této soustavy dostaneme:

$$b_0 = \frac{\sum y_t \cdot \sum t'^4 - \sum t'^2 \cdot \sum y_t \cdot t'^2}{n \sum t'^4 - \left(\sum t'^2 \right)^2}$$

$$b_1 = \frac{\sum y_t \cdot t'}{\sum t'^2}$$

$$b_2 = \frac{n \sum y_t \cdot t'^2 - \sum y_t \cdot \sum t'^2}{n \sum t'^4 - \left(\sum t'^2 \right)^2}$$

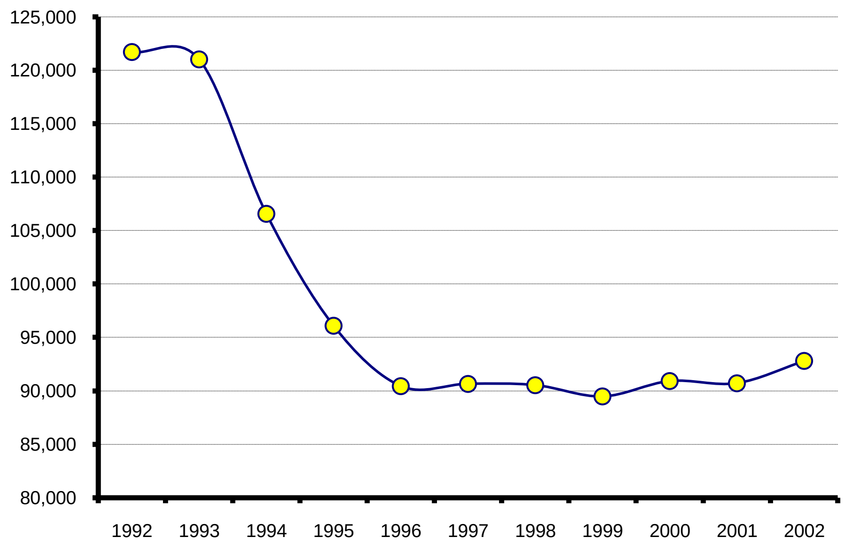
Příklad 2.

Časovou řadu počtu živě narozených dětí v letech 1992 až 2002 vyrovnáme pomocí parabolické trendové funkce a provedeme extrapolaci pro roky 2003 a 2004.

Zdroj: Stránky Českého statistického úřadu

http://www.czso.cz/csu/2006edicniplan.nsf/p/4027-06

Rok	Počet živě narozených dětí
1992	121705
1993	121025
1994	106579
1995	96097
1996	90446
1997	90657
1998	90535
1999	89471
2000	90910
2001	90715
2002	92786
2003	93685
2004	97664



Řešení:

Tabulka 5.

rok	y_t	t'	t'^2	t'^4	$y_t t'$	$y_t t'^2$	Y_t
1992	121705	-5	25	625	-608525	3042625	124300
1993	121025	-4	16	256	-484100	1936400	114607
1994	106579	-3	9	81	-319737	959211	106409
1995	96097	-2	4	16	-192194	384388	99707
1996	90446	-1	1	1	-90446	90446	94501
1997	90657	0	0	0	0	0	90789
1998	90535	1	1	1	90535	90535	88573
1999	89471	2	4	16	178942	357884	87853
2000	90910	3	9	81	272730	818190	88627
2001	90715	4	16	256	362860	1451440	90897
2002	92786	5	25	625	463930	2319650	94663
Celkem	1080926	0	110	1958	-326005	11450769	1080926
Průměr	98266	0	10	178	-29637	1040979	98266

Zdroj: vlastní výpočet

Potřebné hodnoty pro výpočet koeficientů dostaneme z tabulky 5.

$$b_0 = \frac{1080926 \cdot 1958 - 110 \cdot 11450769}{11 \cdot 1958 - 110^2} = 90789,21$$

$$b_1 = \frac{-326005}{110} = -2963,68$$

$$b_2 = \frac{11 \cdot 11450769 - 1080926 \cdot 110}{11 \cdot 1958 - 110^2} = 747,68$$

Hledaná parabolická funkce má tedy tvar:

$$Y_t = 90789,21 - 2963,68 \cdot t' + 747,68 \cdot t'^2$$

Vyrovnané hodnoty parabolickou trendovou funkcí jsou v posledním sloupci tabulky 5.

Nyní vypočteme I^2, použijeme přitom údaje z tabulky 6.

Tabulka 6.

rok	Y_t	Y_t	S_Y	S_T
1992	121705	124300	549386721	677748401,8
1993	121025	114607	517972081	267021881,8
1994	106579	106409	69105969	66314409,3
1995	96097	99707	4704561	2077307,4
1996	90446	94501	61152400	14178489,9
1997	90657	90789	57896881	55902461,6
1998	90535	88573	59768361	93950317,6
1999	89471	87853	77352025	108439744,3
2000	90910	88627	54110736	92905018,5
2001	90715	90897	57017601	54297008,0
2002	92786	94663	30030400	12983171,2
Celkem	**1080926**	**1080926**	**1538497736**	**1445818211**
Průměr	**98266**	**98266**	**139863430,5**	**131438019,2**

Index determinace $I^2 = \Sigma S_T / \Sigma S_Y = \dfrac{1445818211}{1538497736} = 0{,}93976$

Tato hodnota znamená, že 93,98 procent variability hodnot časové řady lze vysvětlit parabolickou trendovou funkcí.

Předpovědi na roky 2003 a 2004:

$$Y_{2003} = 90789{,}21 - 2963{,}68 \cdot 6 + 747{,}68 \cdot 6^2 = 99923{,}61$$

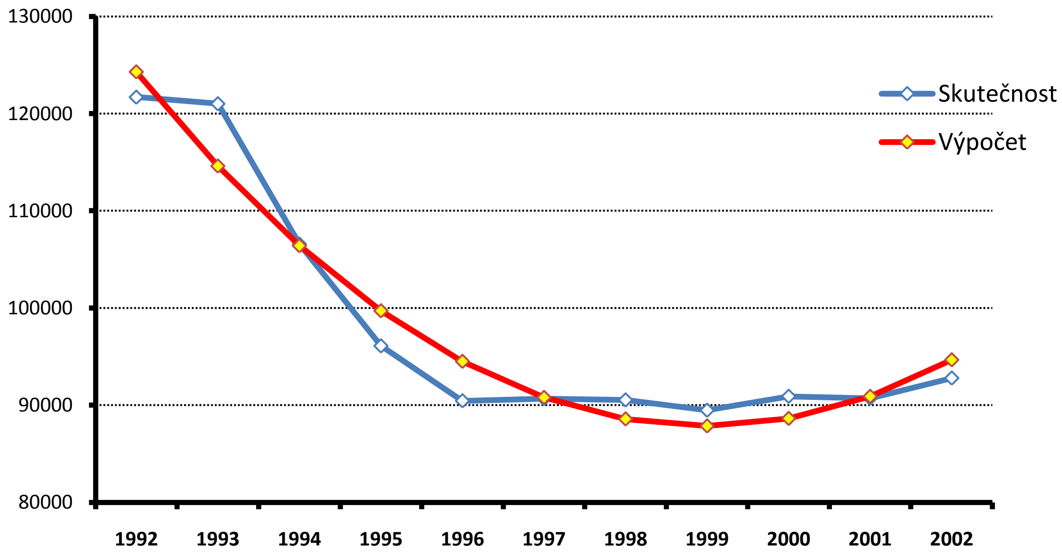
$$Y_{2004} = 90789{,}21 - 2963{,}68 \cdot 7 + 747{,}68 \cdot 7^2 = 106679{,}77$$

Skutečnost:

$y_{2003} = 93685$

$y_{2004} = 97664$

Graf skutečných a vypočtených hodnot časové řady



3.1.3 Exponenciální trend

Exponenciální trendová funkce má tvar:

$Y_t = \beta_0 \cdot \beta_1^t$, kde $t = 1, 2, ..., n$

Tato funkce ale není z hlediska parametrů lineární. Abychom mohli použít metodu nejmenších čtverců, musíme ji nejdříve zlogaritmizovat.

$$\log Y_t = \log(b_0 \cdot b_1^t)$$

$$\log Y_t = \log b_0 + \log b_1^t$$

$$\log Y_t = \log b_0 + t \cdot \log b_1$$

Zde již můžeme použít metodu nejmenších čtverců, tj

$$S(b_0, b_1) = \sum_{t=1}^{n} \left(\log y_t - \log b_0 - t \cdot \log b_1\right)^2$$

Provedeme-li obě parciální derivace a položíme je rovné nule, dostaneme soustavu těchto dvou rovnic:

$$\sum \log y_t = n \cdot \log b_0 + \sum t \cdot \log b_1$$

$$\sum t \cdot \log y_t = \sum t \cdot \log b_0 + \sum t^2 \cdot \log b_1$$

Použijeme-li transformovanou časovou proměnnou *t'*, dostaneme tato řešení pro jednotlivé koeficienty b_0 a b_1:

$$\log b_0 = \frac{\sum \log y_i}{n}$$

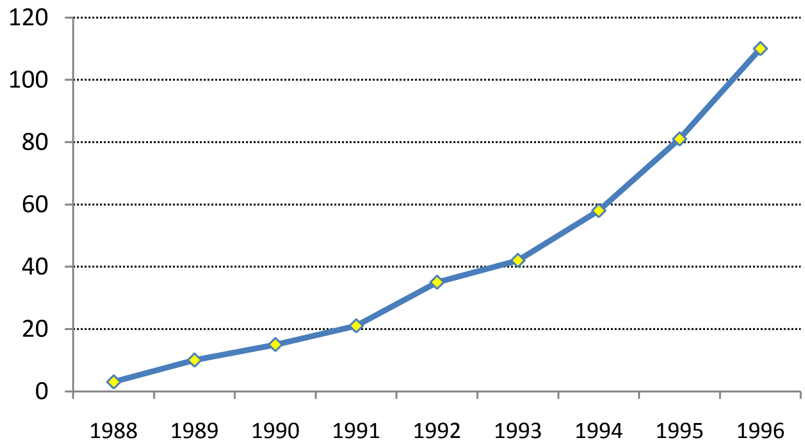
$$\log b_1 = \frac{\sum t' \cdot \log y_i}{\sum (t')^2}$$

Příklad 3.

Následující časovou řadu vyrovnáme pomocí exponenciální trendové funkce:

Zdroj : [2.]

rok	y_i
1988	3
1989	10
1990	15
1991	21
1992	35
1993	42
1994	58
1995	81
1996	110



Řešení:

Tabulka 7.

rok	Y_t	t'	t'^2	$\log y_t$	$\log y_t \cdot t'$	Y_t
1988	3	-4	16	0.477	-1.90849	5.3856
1989	10	-3	9	1	-3	8.0452
1990	15	-2	4	1.176	-2.35218	12.0182
1991	21	-1	1	1.322	-1.32222	17.9533
1992	35	0	0	1.544	0	26.8194
1993	42	1	1	1.623	1.623249	40.0640
1994	58	2	4	1.763	3.526856	59.8492
1995	81	3	9	1.908	5.725455	89.4051
1996	110	4	16	2.041	8.165571	133.557
Součet	**375**	**0**	**60**	**12.856**	**10.4582**	**393.0975**

Potřebné hodnoty pro výpočet koeficientů dostaneme z tabulky 7.

$$\log b_0 = \frac{\sum \log y_t}{n} = \frac{12{,}856}{9} = 1{,}42845$$

$$\log b_1 = \frac{\sum t' \cdot \log y_i}{\sum (t')^2} = \frac{10{,}4582}{60} = 0{,}17431$$

Odlogaritmováním dostaneme hledané koeficienty

$$b_0 = 10^{1,42845} = 26,81946$$

$$b_1 = 10^{0,17431} = 1,49384$$

Hledaná exponenciální funkce má potom tvar:

$$Y_t = 26,81946 \cdot 1,49384^{t'}$$

Vyrovnané hodnoty exponenciální trendovou funkcí jsou v posledním sloupci tabulky 7.

Index determinace:

Tabulka 8.

rok	y_i	Y_i	S_Y	S_T	S_R
1988	3	5.386	1.3057	0.7895	0.0646
1989	10	8.045	0.3841	0.5102	0.0089
1990	15	12.018	0.1969	0.2915	0.0093
1991	21	17.953	0.0885	0.1337	0.0046
1992	35	26.819	0.0057	0.0366	0.0134
1993	42	40.064	0.00001	0.0003	0.0004
1994	58	59.849	0.0206	0.0247	0.0002
1995	81	89.405	0.0833	0.1099	0.0018
1996	110	133.557	0.1777	0.2559	0.0071
Součet	**375**	**393.097**	**2.2627**	**2.1524**	**0.1103**

Protože exponenciální trendová funkce není lineární v parametrech (jak již bylo řečeno výše), musíme i pro index determinace, potažmo čtverce odchylek (S_Y a S_T) použít logaritmy příslušných hodnot, nikoliv hodnoty samotné jako u lineární nebo parabolické trendové funkce.

K výpočtu použijeme tabulku 8. Konkrétně první řádek sloupce S_Y:

$$S_{Y(1988)} = (\log y_1 - \log y)^2 = \left(\log 3 - \log \frac{375}{9}\right)^2 = 1,3057$$

Analogicky pro S_T:

$$S_{T(1988)} = (\log Y_1 - \log y)^2 = \left(\log 5{,}386 - \log \frac{375}{9}\right)^2 = 0{,}7895$$

atd.

$$I^2 = \Sigma S_T / \Sigma S_Y = \frac{2{,}1524}{2{,}2627} = 0{,}95125$$

Což znamená, že přibližně 95% hodnot této časové řady lze vysvětlit pomocí trendové exponenciály.

Kromě lineárního, parabolického a exponenciálního trendu můžeme ještě použít **modifikovaný (posunutý) exponenciální trend** s rovnicí: $Y_t = \kappa + \beta_0 \cdot \beta_1^{\,t}$

a **logistický trend** s rovnicí $Y_t = \dfrac{\kappa}{1 + \beta_0 \cdot \beta_1^{\,t}}$, kde κ je konstanta.

Vždy je třeba použít takový trend, který nejlépe vystihuje empirická data.

Jak zvolit vhodnou trendovou funkci:

1) posouzení grafu
2) první diference (např. lineární trend má přibližně konstantní první diference).
3) index determinace (čím více se blíží číslu 1, tím je vhodnější.
4) interpolační kritéria, z nichž nejpoužívanější je M.S.E. (střední čtvercová chyba odhadu).

$$M.S.E. = \frac{\sum_{t=1}^{n} (y_t - Y_t)^2}{n}$$

Při porovnáváni je nutný stejný počet parametrů.

3.2 Vyrovnávání časových řad pomocí adaptivních metod

Mezi adaptivní metody řadíme metodu klouzavých průměrů, Brownovu metodu exponenciálního vyrovnávání, Box-Jenkinsovu metodu, a další. Zde si ukážeme postup metody klouzavých průměrů a Brownovu metodu exponenciálního vyrovnávání.

3.2.1 Metoda klouzavých průměrů

Podstata klouzavých průměrů

Klouzavé průměry slouží k očištění časové řady od sezónních výkyvů, tj. odstranění sezónní složky. Podstata vyrovnávání pomocí klouzavých průměrů spočívá v nahrazení původních hodnot řadou průměrů, které vypočítáme přímo z těchto hodnot. Počet těchto pozorování, ze kterých budou klouzavé průměry vypočítány nazýváme *klouzavou částí období* a značíme ho m $(=2p+1)$.

Očistit časovou řadu od sezónních výkyvů je možno pomocí

a) prostých klouzavých průměrů
b) centrovaných klouzavých průměrů
c) vážených klouzavých průměrů

Je-li počet hodnot v intervalu sudé číslo (kvartály, měsíce), je celá věc velmi jednoduchá, protože n-členné centrované průměry jsou zároveň váženými klouzavými průměry.

Je-li počet hodnot v intervalu liché číslo (např. dny v týdnu), je celá věc o něco složitější.

3.2.1.1 sudý počet hodnot v intervalu

U sudých musíme centrovat, protože řešením rovnice 2p + 1 = sudé číslo není celé číslo. Tím dostaneme zároveň vážené klouzavé průměry.

Tabulka 9.

1	2	3	4	5	6	7
ROK	kvartál	y_i	4-členné klouzavé úhrny	4-členné klouzavé průměry	4-členné centrované průměry	pomocí vektorů
2000	I.	22 724	-	-	-	
	II.	24 211	90 910	22 727,5	-	
	III.	22 972	90 436	22 609,0	22 668,25	22 668,25
	IV.	21 003	90 250	22 562,5	22 585,75	22 585,75
2001	I.	22 250	90 651	22 662,8	22 612,63	22 612,63
	II.	24 025	90 715	22 678,8	22 670,75	22 670,75
	III.	23 373	91 247	22 811,8	22 745,25	22 745,25
	IV.	21 067	91 618	22 904,5	22 858,13	22 858,13
2002	I.	22 782	92 157	23 039,3	22 971,88	22 971,88
	II.	24 396	92 786	23 196,5	23 117,88	23 117.88
	III.	23 912	-	-	-	
	IV.	21 696	-	-	-	

Zdroj: Zadání semestrální práce ze statistiky

Ve 4. sloupci tabulky 9. Jsou tzv. 4-členné klouzavé úhrny. Tj. součty čtyř příslušných po sobě následujících hodnot. Například první hodnota 90 910 je součtem prvních čtyř hodnot sloupce 3. Druhá hodnota, 90 436 je součtem 2. až 5. atd.

V 5. sloupci jsou průměry těch samých hodnot.

V 6. Sloupci jsou 4-členné centrované průměry. Tj. průměry dvou sousedních hodnot. Například 22 668,25 je průměr prvních dvou hodnot 4-členných centrovaných průměrů. Atd.

Nebo můžeme použít rovnou příslušné vektory, pro 4-členné průměry,

je to vlastně aritmetický průměr dvou sousedních klouzavých průměrů:

$$\frac{1}{8}(1,2,2,2,1)$$

tedy první hodnota = $\frac{1}{8}(22724 + 2 \cdot 24211 + 2 \cdot 22972 + 2 \cdot 21003 + 22250)$ = 22 668,25

druhá hodnota = $\frac{1}{8}(24211 + 2 \cdot 22972 + 2 \cdot 21003 + 2 \cdot 22250 + 24025)$ = 22 585,75

Příslušné hodnoty jsou ve sloupci 7.

Jak je vidět porovnáním sloupců 6 a 7, vedou oba postupy ke stejnému výsledku.

Pro 12-členné klouzavé průměry bychom postupovali obdobně:

Tabulka 10.

ROK	měsíc	y_i	12-členné klouzavé úhrny	12-členné klouzavé průměry	12-členné centrované průměry	pomocí vektorů
2000	1	157	-	-	-	-
	2	172	-	-	-	-
	3	191	-	-	-	-
	4	207	-	-	-	-
	5	220	-	-	-	-
	6	256	2762	230,17	-	-
	7	312	2749	229,08	229,63	229,63
	8	345	2746	228,83	228,96	228,96
	9	301	2744	228,67	228,75	228,75
	10	267	2736	228,00	228,33	228,33
	11	189	2733	227,75	227,88	227,88
	12	145	2743	228,58	228,17	228,17
2001	1	144	2761	230,08	229,33	229,33
	2	169	2780	231,67	230,88	230,88
	3	189	2784	232,00	231,83	231,83
	4	199	2778	231,50	231,75	231,75
	5	217	2757	229,75	230,63	230,63
	6	266	2745	228,75	229,25	229,25
	7	330	-	-	-	-
	8	364	-	-	-	-
	9	305	-	-	-	-
	10	261	-	-	-	-
	11	168	-	-	-	-
	12	133	-	-	-	-

Zdroj: Zadání semestrální práce ze statistiky

Příslušný vektor pro 12-členné průměry je : $\frac{1}{24}(1,2,2,2,2,2,2,2,2,2,2,2,1)$

Tímto jsme s průměrováním hotovi, protože prosté **centrované** klouzavé průměry jsou zároveň **vážené** klouzavé průměry.

3.2.1.2 lichý počet hodnot v intervalu

a) prosté

Často používané jsou 7-členné klouzavé průměry (např. u denních hodnot).

U 7-členných se m = 7, proto 3 řádky ($m = 2p + 1$) musíme vynechat, centrování není nutné.

Tabulka 11.

týden	**den**	**y_i**	**7-členné klouzavé úhrny**	**7-členné prosté klouzavé průměry**
1	Po	6 125	-	-
	Út	6 657	-	-
	St	5 799	-	-
	Čt	4 691	46 984	6 712,00
	Pá	8 600	46 846	6 692,29
	So	9 111	46 389	6 627,00
	Ne	6 001	46 702	6 671,71
2	Po	5 987	47 111	6 730,14
	Út	6 200	46 733	6 676,14
	St	6 112	47 383	6 769,00
	Čt	5 100	46 882	6 697,43
	Pá	8 222	46 907	6 701,00
	So	9 761	46 930	6 704,29
	Ne	5 500	46 710	6 672,86
3	Po	6 012	46 110	6 587,14
	Út	6 223	46 653	6 664,71
	St	5 892	46 447	6 635,29
	Čt	4 500	46 943	6 706,14
	Pá	8 765	-	-
	So	9 555	-	-
	Ne	5 996	-	-

Zdroj: [7]

Postup je úplně stejný jako u sudých klouzavých průměrů.

b) vážené

Pro vážené 7-členné klouzavé průměry použijeme vektor: $\frac{1}{21}(-2,3,6,7,6,3,-2)$

A to proto, že váhy $W_i = \frac{3}{4m(m^2-4)}\left[(3m^2-7)-20i^2\right]$

Výpočet vah pro 7-členné vážené klouzavé průměry:

Tabulka 12.

	i	**i²**	**20.i²**	**140 - 20.i²**	**W_i**	**W_i`**
	-3	9	180	-40	-0.09524	-2
	-2	4	80	60	0.142857	3
	-1	1	20	120	0.285714	6
	0	0	0	140	0.333333	7
	1	1	20	120	0.285714	6
	2	4	80	60	0.142857	3
	3	9	180	-40	-0.09524	-2
Σ	**0**				**1**	**21**

Zdroj : [2.]

Výsledný vektor vah je v posledním sloupci tabulky 12.

Potom jednotlivé 7-členné vážené klouzavé průměry jsou v posledním sloupci tabulky 13.

Tabulka 13.

týden	den	y_i	7-členné vážené klouzavé průměry
1	Po	6 125	-
	Út	6 657	-
	St	5 799	-
	Čt	4 691	6 775,38
	Pá	8 600	7 291,62
	So	9 111	7 591,38
	Ne	6 001	7 399,48
2	Po	5 987	6 351,62
	Út	6 200	5 458,62
	St	6 112	5 794,62
	Čt	5 100	6 981,57
	Pá	8 222	7 482,48
	So	9 761	7 586,91
	Ne	5 500	7 356,62
3	Po	6 012	6 377,95
	Út	6 223	5 139,67
	St	5 892	5 704,91
	Čt	4 500	6 798,09
	Pá	8 765	-
	So	9 555	-
	Ne	5 996	-

Zdroj: [7]

Například první hodnotu (6 775,38) dostaneme pomocí vypočteného vektoru:

$$\frac{1}{21}\left(-2\cdot 6125+3\cdot 6657+6\cdot 5799+7\cdot 4691+6\cdot 8600+3\cdot 9111-2\cdot 6001\right)=6775{,}38$$

Podobně bychom udělali 5-členné vážené průměry:

Tabulka 14.

	i	**i^2**	**$20.i^2$**	**$68 - 20.i^2$**	**W_i**	**W_i`**
	-2	4	80	-12	-0.085714	-3
	-1	1	20	48	0.3428571	12
	0	0	0	68	0.4857143	17
	1	1	20	48	0.3428571	12
	2	4	80	-12	-0.085714	-3
Σ	**0**				**1**	**35**

Zdroj : [2.]

Tedy dostaneme vektor $\frac{1}{35}(-3,12,17,12,-3)$

Stejným způsobem bychom dostali pro 9-členné vážené průměry vektor

$$\frac{1}{231}(-21,14,39,54,59,54,39,14,-21)$$

Nevýhodou této metody je, že při použití intervalu délky *2p+1* ztratíme *p* hodnot na začátku i na konci každé řady.

3.2.2 Exponenciální vyrovnávání (Brownova metoda)

Na rozdíl od klasických metod vyrovnávání časových řad (např. metodou nejmenších čtverců), kde je každé hodnotě časové řady přisuzována stejná váha. U exponenciálního vyrovnávání adaptivní metodou je novějším pozorováním přisuzována větší váha než pozorováním starším. Dalším významným rozdílem je, že klasické metody při odhadování parametrů nerozlišují mezi staršími a novějšími pozorováními. Zatímco u adaptivních metod vzorce pro extrapolaci průběžně upravujeme pomocí nově získaných empirických údajů.

Hodnotu trendové složky T_{n-k} lze popsat funkcí :

$$T_{n-k} = a_0 - a_1 k + a_2 k^2 + \ldots + (-1)^k a_k k^k$$

U adaptivní metody minimalizujeme součet:

$$S(a_1, a_2, \ldots, a_n) = \sum_{k=0}^{n-1} (y_{n-k} - T_{n-k})^2 \cdot w_k \rightarrow MIN.$$

kde w_k jsou váhy, které jsou nepřímo úměrné stáří pozorování.

$w_k = \alpha^k$, kde $k = 0,1, \ldots , n\text{-}1$ a α je *vyrovnávací konstanta* a platí $0 < \alpha < 1$.

Váha w_k je exponenciální funkcí stáří pozorovaných hodnot.

Známe tři druhy exponenciálního vyrovnávání:

1) Jednoduché (trend časové řady je zhruba konstantní)
2) Dvojité (trend roste nebo klesá lineárně)
3) Trojité (trend je přibližně kvadratický)

V této práci si ukážeme postup výpočtu jednoduchého a dvojitého vyrovnávání.

3.2.2.1 Jednoduché exponenciální vyrovnávání:

Podle [1,2]: $k = 0$ proto $T_{n-k} = a_0$

$$\sum_{k=0}^{n-1}(y_{n-k} - a_0)^2 \cdot \alpha^k \to MIN.$$

$$\frac{dS}{da_0} = \sum_{k=0}^{n-1} 2(y_{n-k} - a_0)\cdot(-1)\cdot\alpha^k = -2\cdot\sum_{k=0}^{n-1}(y_{n-k} - a_0)\cdot\alpha^k$$

derivaci položíme rovnou nule:

$$-2\cdot\sum_{k=0}^{n-1}(y_{n-k} - a_0)\cdot\alpha^k = 0$$

$$-\sum_{k=0}^{n-1}(y_{n-k} - a_0)\cdot\alpha^k = 0$$

$$\sum_{k=0}^{n-1}\alpha^k \cdot y_{n-k} = a_0 \cdot \sum_{k=0}^{n-1}\alpha^k$$

α^k je geometrická posloupnost s prvním členem rovným 1 a kvocientem q = α .

Proto pro $\alpha \in (-1;1)$ existuje součet $\sum_{k=0}^{\infty}\alpha^k$ a je roven $\frac{1}{1-\alpha}$.

Potom můžeme při dostatečně vysokém *n* výraz $\sum_{k=0}^{n-1}\alpha^k$ aproximovat výrazem $\sum_{k=0}^{\infty}\alpha^k$.

Pak dostáváme:

$$\sum_{k=0}^{n-1}\alpha^k \cdot y_{n-k} = a_0 \cdot \frac{1}{1-\alpha}$$

odtud

$$a_0 = (1-\alpha)\cdot\sum_{k=0}^{n-1}\alpha^k \cdot y_{n-k}$$

Tento výraz můžeme přepsat do výpočetně výhodnějšího tvaru, který umožňuje počítat vyrovnané hodnoty rekurentně.

$$\hat{y}_t = (1-\alpha)\cdot y_t + \alpha\cdot\hat{y}_{t-1} \qquad (10.)$$

Celý postup má 2 fáze.

V první fázi určíme hodnotu vyrovnávací konstanty α, obvykle pomocí statistického softwaru. Nejprve určíme (podle [1]) počáteční hodnotu $\hat{y}_0$ jako aritmetický průměr prvních 6 hodnot. Další hodnoty $\hat{y}_t$ pak určíme podle vzorce (1.).

Nakonec zjistíme střední chybu odhadu (SSE) pro jednotlivé hodnoty vyrovnávací konstanty α (pro α=0,50 až 0,99).

Ve druhé fázi již pokračujeme s tou konkrétní hodnotou α, pro kterou byla SSE nejmenší. Opět určíme aritmetický průměr, tentokrát všech hodnot časové řady, které máme k dispozici. Zbytek výpočtu je stejný jako v první fázi.

3.2.2.2 Dvojité exponenciální vyrovnávání:

Podle [1,2]: $k = 1$ proto $T_{n-k} = a_0 + a_1 \cdot t$

Budeme minimalizovat součet:

$$\sum_{k=0}^{n-1} \left(y_{n-k} - (a_0 + a_1 \cdot k)\right)^2 \cdot \alpha^k \to MIN.$$

Celý postup má opět dvě fáze. V první fázi určíme hodnotu vyrovnávací konstanty α.

Nejprve určíme koeficienty b_0 a b_1 pomocí metody nejmenších čtverců lineární trendové funkce prvních 6 hodnot časové řady.

Potom budeme potřebovat dvě nové proměnné, a to

$\boldsymbol{S_t}$ jako jednoduchou vyrovnávací statistiku a

$\boldsymbol{S_t'}$ jako dvojitou vyrovnávací statistiku.

Určíme počáteční hodnoty vyrovnávacích statistik pomocí vzorců

$$S_0 = b_0 - \frac{\alpha}{1-\alpha} \cdot b_1 \qquad (11.)$$

$$S_0' = b_0 - \frac{2\alpha}{1-\alpha} \cdot b_1 \qquad (12.)$$

Dále určíme hodnoty vyrovnávacích statistik pro všechny hodnoty řady.

$$S_t = (1-\alpha) y_t + \alpha \cdot S_{t-1} \qquad (13.)$$

$$S_t' = (1-\alpha) S_t + \alpha \cdot S_{t-1}' \qquad (14.)$$

Dále určíme hodnoty Y_t pomocí vzorce

$$Y_t = \left(2 + \frac{(1-\alpha)}{\alpha}\right) \cdot S_t - \left(1 + \frac{(1-\alpha)}{\alpha}\right) \cdot S_t' \qquad (15.)$$

A nakonec určíme součet čtverců odchylek y_t od Y_t, tzv. součet čtvercových chyb SSE. Celý tento postup opakujeme pro α = 0,50 až 0,99. Pro druhou fázi výpočtu použijeme tu hodnotu vyrovnávací konstanty α, pro kterou nám SSE vyšlo nejmenší.

Ve druhé fázi, podobně jako v první, opět určíme koeficienty b_0 a b_1 lineární trendové funkce pomocí metody nejmenších čtverců. Tentokrát ale nepoužijeme prvních 6 hodnot, ale všechny hodnoty časové řady.

Dále určíme počáteční hodnoty vyrovnávacích statistik pomocí vzorců (11.) a (12.).

Z nich určíme jednotlivé vyrovnávací statistiky S_t a S'_t podle vzorců (13.) a (14.).

Potom určíme jednotlivé vyrovnané hodnoty Y_t pomocí vzorce

$$Y_t = 2S_t - S'_t$$

Předpovědi určíme podle vzorce

$$Y_{n+t} = \left(2 + \frac{(1-\alpha)\cdot t}{\alpha}\right)\cdot S_n - \left(1 + \frac{(1-\alpha)\cdot t}{\alpha}\right)\cdot S'_n$$, kde t = 1, 2, 3, - následujících let. (16.)

3.4 Souhrnný příklad

Následující časovou řadu vyrovnáme nejprve pomocí lineární trendové funkce (klasickou metodou) a poté pomocí adaptivní metody dvojitého exponenciálního vyrovnávání.

Počet rozvodů v letech 1958 – 1981
http://www.czso.cz/csu/edicniplan.nsf/t/93003E0F91/$File/tab1.xls

Tabulka zadání:

rok	počet rozvodů	rok	počet rozvodů	rok	počet rozvodů	rok	počet rozvodů
1958	13 589	1964	14 446	1970	21 516	1976	25 544
1959	13 222	1965	16 196	1971	23 616	1977	25 442
1960	12 970	1966	17 435	1972	22 392	1978	27 071
1961	13 939	1967	17 352	1973	25 271	1979	26 191
1962	14 137	1968	18 647	1974	24 970	1980	27 218
1963	14 703	1969	20 550	1975	26 154	1981	27 608

a) Lineární trend

Všechny hodnoty pro výpočet najdeme v posledních dvou řádcích tabulky 15.

Tabulka 15.

rok	y_t	t'	y_t . t'	t'^2	Y_t
1958	13 589	-23	-312 547	529	11 899,21
1959	13 222	-21	-277 662	441	12 640,51
1960	12 970	-19	-246 430	361	13 381,80
1961	13 939	-17	-236 963	289	14 123,10
1962	14 137	-15	-212 055	225	14 864,40
1963	14 703	-13	-191 139	169	15 605,69
1964	14 446	-11	-158 906	121	16 346,99
1965	16 196	-9	-145 764	81	17 088,29
1966	17 435	-7	-122 045	49	17 829,59
1967	17 352	-5	-86 760	25	18 570,88
1968	18 647	-3	-55 941	9	19 312,18
1969	20 550	-1	-20 550	1	20 053,48
1970	21 516	1	21 516	1	20 794,77
1971	23 616	3	70 848	9	21 536,07
1972	22 392	5	111 960	25	22 277,37
1973	25 271	7	176 897	49	23 018,66
1974	24 970	9	224 730	81	23 759,96
1975	26 154	11	287 694	121	24 501,26
1976	25 544	13	332 072	169	25 242,56
1977	25 442	15	381 630	225	25 983,85
1978	27 071	17	460 207	289	26 725,15
1979	26 191	19	497 629	361	27 466,45
1980	27 218	21	571 578	441	28 207,74
1981	27 608	23	634 984	529	28 949,04
součet	**490 179**	**0**	**1 704 983**	**4 600**	**490 179**
průměr	**20 424,13**	**0**	**71040,96**	**191,67**	**20 424,13**

Pomocí vzorců (8.) a (9.) určíme b_0 a b_1.

$$b_0 = \frac{\sum y_t}{n} = y = 20424{,}13$$

$$b_1 = \frac{\sum y_t t'}{\sum t'^2} = \frac{1704983}{4600} = 370{,}65$$

Hledaná lineární funkce má tvar:

$$Y = 20424{,}13 + 370{,}65 \cdot t'$$

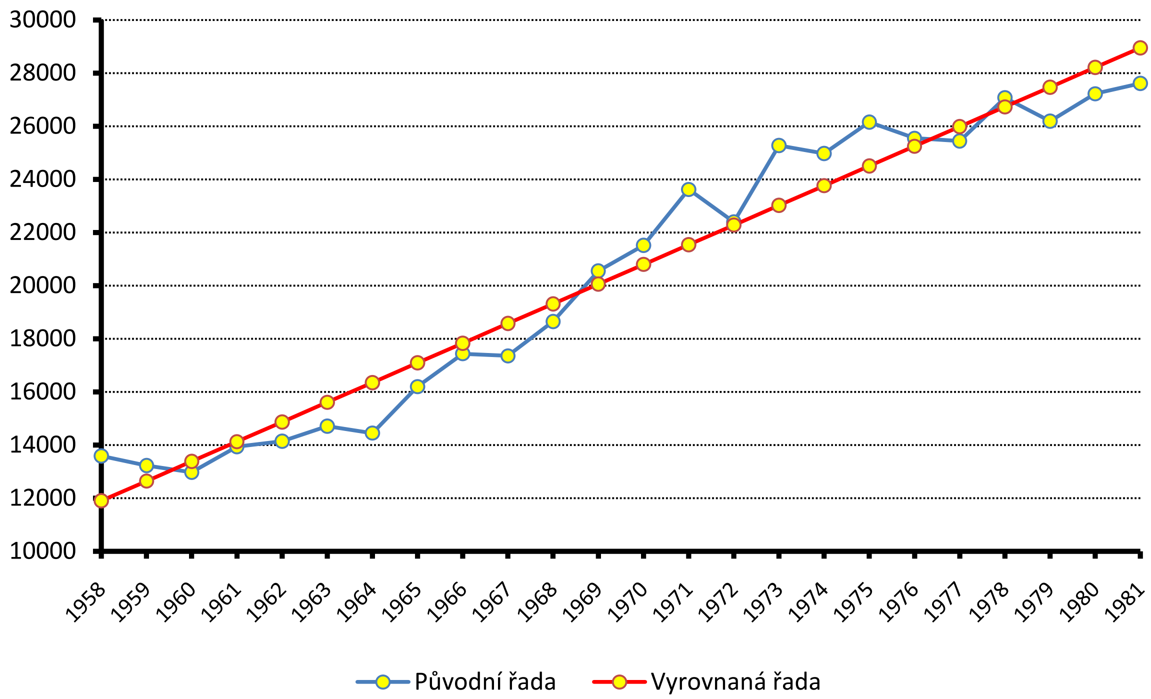
Jednotlivé vyrovnané hodnoty časové řady jsou v šestém sloupci tabulky 15.

Pro doplnění index determinace je:

$$I^2 = \frac{S_T}{S_Y} = \frac{631949354}{662467649} = 0{,}954$$

To znamená, že asi 95% variability hodnot dané časové řady lze vysvětlit pomocí lineární trendové funkce.

Graf skutečných hodnot a hodnot vyrovnaných lineární trendovou funkcí:



Předpovědi na další tři roky:	Skutečné hodnoty:	Rozdíl:
$Y_{1982} = 20424{,}13 + 370{,}65 \cdot 25 = 29690{,}38$	27 821	1869,38
$Y_{1982} = 20424{,}13 + 370{,}65 \cdot 27 = 30431{,}68$	29 319	1112,68
$Y_{1982} = 20424{,}13 + 370{,}65 \cdot 29 = 31172{,}98$	30 514	658,98

b) dvojité exponenciální vyrovnávání (Brownova metoda)

1. fáze:

Nejdříve určíme vyrovnávací konstantu α.

Vyrovnáme prvních šest hodnot časové řady lineární trendovou funkcí, abychom určili koeficienty b_0 a b_1.

Tabulka 16.

t	y_t	$t \cdot y_t$	t^2	Y_t
1	13 589	13 589	1	13 096,86
2	13 222	26 444	4	13 362,11
3	12 970	38 910	9	13 627,37
4	13 939	55 756	16	13 892,63
5	14 137	70 685	25	14 157,89
6	14 703	88 218	36	14 423,14
21	**82 560**	**293 602**	**91**	**82 560**
3.5	**13 760**	**48 933,67**	**15,167**	**13 760**

Zdroj: vlastní výpočet

Pomocí vzorců (3.) a (4.) a příslušných hodnot z tabulky 16 vypočteme parametry b_0 a b_1.

$$b_1 = \frac{48933,67 - 13760 \cdot 3,5}{15,167 - 3,5^2} = 265,23$$

$$b_0 = 13760 - 265,23 \cdot 3,5 = 12831,69$$

S těmito hodnotami určíme hodnotu vyrovnávací konstanty α pomocí statistického softwaru nebo v Excelu.

Jako optimální hodnotu vyrovnávací konstanty α použijeme α = 0,59.

Určíme počáteční hodnoty vyrovnávacích statistik pomocí vzorců (11.) a (12.).

Pro α = 0,59

$$S_0 = b_0 - \frac{\alpha}{1-\alpha} \cdot b_1 = 12831,69 - \frac{0,59}{0,41} \cdot 265,23 = 12449,89$$

$$S_0' = b_0 - \frac{2\alpha}{1-\alpha} \cdot b_1 = 12831,69 - \frac{1,18}{0,41} \cdot 265,23 = 12068,18$$

Pomocí vzorců (13.), (14.) a (15.) vypočteme hodnoty ve třetím, čtvrtém a pátém sloupci tabulky 17.

Tabulka 17.

t	y_t	S_t	S_t'	Y_t	SSE
1	13 589	12916.92	12416.16	13096.86	242204.59
2	13 222	13042.01	12672.76	13765.67	295578.62
3	12 970	13012.48	12812.05	13667.85	486990.73
4	13 939	13392.36	13049.97	13352.21	344325.22
5	14 137	13697.66	13315.52	13972.66	27006.17
6	14 703	14109.85	13641.20	14345.35	127916.05
7	14 446	14247.67	13889.85	14904.17	209923.44
8	16 196	15046.49	14364.07	14854.14	1800576.05
9	17 435	16025.78	15045.37	16203.12	1517527.87
10	17 352	16569.53	15670.28	17687.48	112547.98
11	18 647	17421.29	16388.19	18093.69	306156.63
12	20 550	18704.06	17337.70	19172.31	1898035.16
13	21 516	19856.96	18370.59	21019.93	246083.30
14	23 616	21398.16	19611.90	22376.21	1537068.03
15	22 392	21805.64	20511.33	24425.73	4136075.41
16	25 271	23226.44	21624.52	23999.38	1617027.95
17	24 970	23941.3	22574.40	25941.54	943891.11
18	26 154	24848.51	23506.78	26258.07	10830.62
19	25 544	25133.66	24173.80	27122.61	2492008.34
20	25 442	25260.08	24619.18	26760.53	1738527.95
21	27 071	26002.56	25186.36	26346.35	525111.33
22	26 191	26079.82	25552.68	27385.94	1427874.33
23	27 218	26546.47	25960.13	26973.27	59890.44
24	27 608	26981.7	26378.98	27540.27	4587.80
Součet					**22107765.14**

V posledním sloupci tabulky 17. jsou hodnoty střední chyby odhadu (SSE).

Součet SSE (22107765,14) je minimální právě pro $\alpha = 0,59$.

Druhá fáze:

Všechny hodnoty vyrovnáme lineární trendovou funkcí

Tabulka 18.

	t	y_t	$t \cdot y_t$	t^2	***Yt***
	1	13589	13589	1	11899.21
	2	13222	26444	4	12640.51
	3	12970	38910	9	13381.80
	4	13939	55756	16	14123.10
	5	14137	70685	25	14864.40
	6	14703	88218	36	15605.69
	7	14446	101122	49	16346.99
	8	16196	129568	64	17088.29
	9	17435	156915	81	17829.59
	10	17352	173520	100	18570.88
	11	18647	205117	121	19312.18
	12	20550	246600	144	20053.48
	13	21516	279708	169	20794.77
	14	23616	330624	196	21536.07
	15	22392	335880	225	22277.37
	16	25271	404336	256	23018.66
	17	24970	424490	289	23759.96
	18	26154	470772	324	24501.26
	19	25544	485336	361	25242.56
	20	25442	508840	400	25983.85
	21	27071	568491	441	26725.15
	22	26191	576202	484	27466.45
	23	27218	626014	529	28207.74
	24	27608	662592	576	28949.04
součet	**300**	**490179**	**6979729**	**4900**	**490179**
průměr	**12.5**	**20424.13**	**290822**	**204.1667**	**20424.13**

$$b_1 = \frac{290822 - 20424{,}13 \cdot 12{,}5}{204{,}167 - 12{,}5^2} = 741{,}3$$

$$b_0 = 20424{,}13 - 741{,}3 \cdot 12{,}5 = 11157{,}9$$

Podle vzorců (11.) a (12.):

$$S_0 = b_0 - \frac{\alpha}{1-\alpha} \cdot b_1 = 11157{,}9 - \frac{0{,}59}{0{,}41} \cdot 741{,}3 = 10091{,}15$$

$$S_0' = b_0 - \frac{2\alpha}{1-\alpha} \cdot b_1 = 11157{,}9 - \frac{2 \cdot 0{,}59}{0{,}41} \cdot 741{,}3 = 9024{,}4$$

Tabulka 19.

t	y_t	S_t	S_t'	Y_t
1	13589	11525.28	10049.77	13000.78
2	13222	12220.93	10939.95	13501.92
3	12970	12528.05	11591.07	13465.03
4	13939	13106.54	12212.41	14000.67
5	14137	13529.03	12752.23	14305.83
6	14703	14010.36	13268.06	14752.65
7	14446	14188.97	13645.63	14732.31
8	16196	15011.85	14205.78	15817.92
9	17435	16005.34	14943.60	17067.08
10	17352	16557.47	15605.29	17509.66
11	18647	17414.18	16346.93	18481.42
12	20550	18699.87	17311.64	20088.10
13	21516	19854.48	18354.20	21354.76
14	23616	21396.70	19601.63	23191.78
15	22392	21804.78	20504.92	23104.63
16	25271	23225.93	21620.53	24831.32
17	24970	23941.00	22571.92	25310.07
18	26154	24848.33	23505.25	26191.41
19	25544	25133.55	24172.85	26094.25
20	25442	25260.02	24618.59	25901.44
21	27071	26002.52	25186.00	26819.04
22	26191	26079.80	25552.46	26607.14
23	27218	26546.46	25960.00	27132.92
24	27608	26981.69	26378.89	27584.49

Hodnoty S_t , S'_t a Y_t jsou (první řádek sloupců 3, 4 a 5 tabulky 19.):

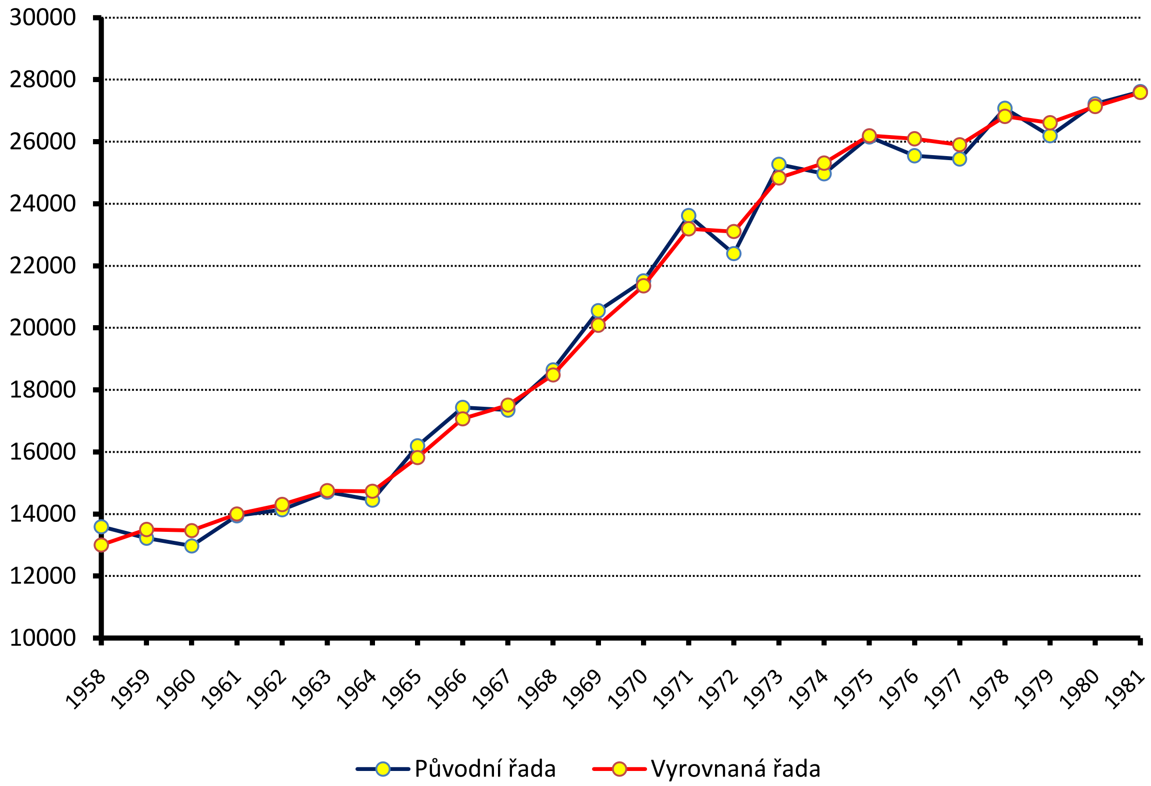
$$S_1 = (1-0{,}59)\cdot 135{,}89 + 0{,}59 \cdot 10091{,}15 = 11525{,}28$$

$$S_1' = (1-0{,}59)\cdot 11525{,}28 + 0{,}59 \cdot 9024{,}4 = 10049{,}77$$

$$Y_1 = 2S_1 - S_1' = 2\cdot 11525{,}28 - 10049{,}77 = 13000{,}78$$

Všechny hodnoty jsou ve sloupcích 3, 4 a 5 v tabulce 19.

Graf skutečných hodnot a vyrovnaných hodnot pomocí adaptivní metody:



Jak můžeme pozorovat z obou grafů, adaptivní metoda exponenciálního vyrovnávání téměř kopíruje empirické hodnoty časové řady.

Předpovědi na další tři roky:

Podle vzorce (16.)

$$Y_{1982}=\left(2+\frac{(1-0{,}59)\cdot 1}{0{,}59}\right)\cdot 26981{,}69-\left(1+\frac{(1-0{,}59)\cdot 1}{0{,}59}\right)\cdot 26378{,}89=28003{,}384$$

$$Y_{1983}=\left(2+\frac{(1-0{,}59)\cdot 2}{0{,}59}\right)\cdot 26981{,}69-\left(1+\frac{(1-0{,}59)\cdot 2}{0{,}59}\right)\cdot 26378{,}89=28423{,}489$$

$$Y_{1984}=\left(2+\frac{(1-0{,}59)\cdot 3}{0{,}59}\right)\cdot 26981{,}69-\left(1+\frac{(1-0{,}59)\cdot 3}{0{,}59}\right)\cdot 26378{,}89=28842{,}736$$

Adaptivní metoda (na rozdíl od klasické metody) může navíc tyto předpovědi průběžně aktualizovat.

Dostaneme-li k dispozici skutečnou hodnotu pro další rok, tj. y_{25} = 27 821, spočítáme nové hodnoty vyrovnávacích statistik podle vzorců (5.) a (6.) na straně 50.

$$S_{25} = (1 - 0{,}59) \cdot 27821 + 0{,}59 \cdot 26981{,}69 = 27325{,}81$$

$$S'_{25} = (1 - 0{,}59) \cdot 27325{,}81 + 0{,}59 \cdot 26378{,}89 = 26767{,}13$$

Předpověď na další rok, tj. 1983 bude:

$$Y_{1983} = \left(2 + \frac{(1 - 0{,}59)}{0{,}59} \right) \cdot 27325{,}81 - \left(1 + \frac{(1 - 0{,}59)}{0{,}59} \right) \cdot 26767{,}13 = 28273$$

4. Závěr

Na závěr je třeba poznamenat, že jakékoli výsledky analýzy je třeba brát s určitou rezervou. I ten nejlepší model nedokáže zahrnout náhodnou složku časové řady a i ty nejpřesnější odhady budoucích hodnot se mohou značně rozcházet s pozdějšími empirickými hodnotami, pokud se změní trend

Vypracujeme-li například analýzu vývoje ceny benzínu a v Saudské Arábii vypukne občanská válka, pak bude cena za litr Naturalu 95 50 korun a všechny výpočty budou bezcenné.

Jinak řečeno, všechny odhady budoucích hodnot jsou relevantní za předpokladu, že se trend vývoje nezmění.

5. Použitá literatura

[1] CIPRA, T.: *Analýza časových řad s aplikacemi v ekonomice*. 1. vyd. Praha: *SNTL*, 1986. 426 s.

[2] HINDLS, R.; HRONOVÁ, S.; NOVÁK, I.: *Metody statistické analýzy pro ekonomy*. 2. vyd. Praha: Management Press, 2000. 259 s. ISBN 80-7261-013-9.

[3] KOZÁK, J., SEGER, J.: *Jednoduché statistické metody v prognostice*. 1. vyd. Praha: *SNTL*, 1975. 278 s.

[4] MAREK, L.: *Statistika pro eko*nomy. 2. vyd. Praha: Professional Publishing, 2007. 485 s. ISBN 978-80-86946-40-5

[5] ARLT, J.; ŠKUTHANOVÁ, M.: *Úvod do problematiky sezónního očišťováni ekonomických časových řad*. 1. vyd. VŠE Praha, 1995. xxx s. ISSN 0572-3043

[6] Statistická ročenka České republiky. Praha: Český statistický úřad. Dostupný na http://www.czso.cz

[7] Přednášky Ing. Vladimíry Hovorkové-Valentové, PhD.

Příloha: Tabulka kvantilů rozdělení t.

v	P				
	0,90	0,95	0,975	0,99	0,995
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771
28	1,313	1,701	2,048	2,467	2,763
29	1,311	1,699	2,045	2,462	2,756
30	1,310	1,697	2,042	2,457	2,750

Pro P < 0,5 jsou hodnoty kvantilů dány vztahem $t_P = -t_{1-P}$

v.......počet stupňů volnosti

$P = 1 - \alpha/2$